

Spontaneous CP violation and quark mass ambiguities

Michael Creutz

Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA

Abstract. I explore the regions of quark masses where CP will be spontaneously broken in the strong interactions. The boundaries of these regions are controlled by the chiral anomaly, which manifests itself in ambiguities in the definition of non-degenerate quark masses. In particular, the concept of a single massless quark is ill defined.

INTRODUCTION

In this talk I discuss two apparently distinct but deeply entwined topics. First, I ask for what quark masses is CP spontaneously broken. Second, I investigate whether a vanishing up quark mass is a physically meaningful concept. I should emphasize at the outset that I will be exploring rather unphysical regions of parameter space. This in some sense is a theorists fantasy, with no direct experimental relevance. Indeed, my real goal is to understand better how chiral symmetry works in the strong interactions. The main content of this talk is contained in two recent papers [1, 2]. The basic ideas have roots in a talk I gave in Como at the 1996 edition of this conference [3].

I require a few assumptions. First, the continuum limit of QCD should exist and confine, with the only relevant parameters being the coupling and the quark masses. Then I will assume that chiral symmetry is spontaneously broken in the usual way and that effective chiral Lagrangians are qualitatively correct. Finally I assume that the anomaly removes any flavor singlet chiral symmetry. In particular, this implies that a single massless quark gives no exact Goldstone boson.

The underlying concepts are all quite old. In 1971 Dashen [4] showed how CP symmetry could be spontaneously broken in the strong interactions. DiVecchia and Veneziano [5] observed the CP violation in chiral Lagrangians at negative quark masses. Georgi and McArthur [6] showed that non-perturbative effects could give a non-multiplicative shift to the up quark mass. This motivated Kaplan and Manohar [7] in their classic studies of ambiguities in the up quark mass in the context of effective chiral Lagrangians. Banks, Nir and Seiberg [8] discussed the fact that the concept of a vanishing up quark mass is not so clean.

THE EFFECTIVE MESON THEORY

I base my initial discussion on the effective theory for pseudoscalar mesons in terms of a field taking values in the group $SU(3)$

$$\Sigma = \exp(i\pi_\alpha \lambda_\alpha / f_\pi) \in SU(3) \quad (1)$$

I work with the three flavor theory, i.e. I include the up, down, and strange quarks. The standard generators of the group $SU(3)$ are given by λ_α , and the pseudoscalar octet fields are denoted as π_α .

Chiral symmetry is manifested in independent left or right global rotations on this field

$$\Sigma \rightarrow g_L^\dagger \Sigma g_R \quad (2)$$

This symmetry is explicitly broken by the quark masses. The lowest order effective Lagrangian including the masses takes the form

$$L = \frac{f_\pi^2}{4} \text{Tr}(\partial_\mu \Sigma^\dagger \partial_\mu \Sigma) - v \text{Re Tr}(\Sigma M) \quad (3)$$

with the mass matrix

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} \quad (4)$$

Expanding this density to quadratic order in meson fields and then diagonalizing the resulting non-derivative term gives the usual result that the meson masses squared are proportional to the quark masses, including $m_{\pi^\pm}^2 \sim m_u + m_d$. For my purposes, I am particularly interested in the isospin violation arising from the up-down mass difference $m_d - m_u$. This results in a mixing of the π^0 and η mesons, giving somewhat complicated formulae for their masses

$$\begin{aligned} m_{\pi^0}^2 &\sim \frac{2}{3} \left(m_u + m_d + m_s - \sqrt{m_u^2 + m_d^2 + m_s^2 - m_u m_d - m_u m_s - m_d m_s} \right) \\ m_\eta^2 &\sim \frac{2}{3} \left(m_u + m_d + m_s + \sqrt{m_u^2 + m_d^2 + m_s^2 - m_u m_d - m_u m_s - m_d m_s} \right) \end{aligned} \quad (5)$$

Note in particular that it is possible to tune the parameters such that $m_{\pi^0}^2$ vanishes. This occurs when

$$m_u = \frac{-m_s m_d}{m_s + m_d}. \quad (6)$$

This vanishing mass does not require chiral symmetry. It does, however, occur at a somewhat unphysical location, requiring at least one of the quark masses to be negative.

SPONTANEOUS CP VIOLATION

Going to negative quark masses at first seems a bit strange, but in such a regime unusual things do happen. Note that because of the anomaly, the signs of the quark masses can

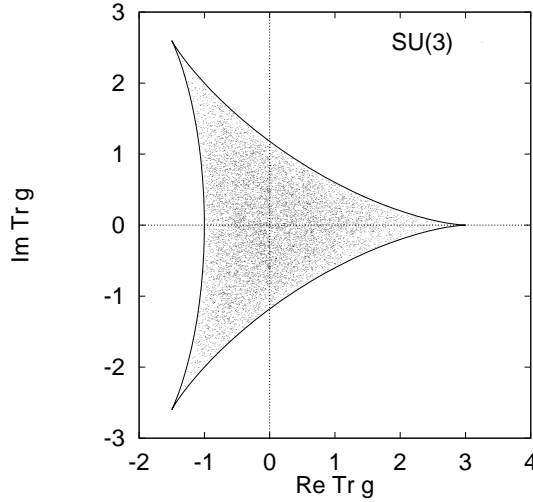


FIGURE 1. The real and imaginary parts of the traces of 10,000 randomly chosen $SU(3)$ matrices. Note that the minimum real part occurs at two distinct cube roots of unity.

become significant. Flavored chiral rotations can move the signs around, but the overall sign of the determinant of the mass matrix is invariant.

The effective Lagrangian is useful for clarifying the expected behavior with negative masses. In the usual case with positive masses, the vacuum involves quantum fluctuations about the maximum of $\text{ReTr}\Sigma$. This occurs at $\Sigma = I$. However, now consider the case of degenerate negative masses. The vacuum instead should occur at the minimum of $\text{ReTr}\Sigma$. The important point is that $-I$ does not lie in the group $SU(3)$. A simple analysis shows that the minimum is doubly degenerate, occurring at $\Sigma = \exp(\pm 2\pi i/3)$. Fig. 1 plots the traces of 10,000 random matrices to illustrate this result. Note that CP symmetry takes Σ to Σ^* ; thus, either of these solutions involves a spontaneous breaking of CP.

This CP violating phase can be approached continuously by passing through the values of the quark masses in Eq. 6 where $m_{\pi^0}^2$ vanishes. Indeed, this equation represents the boundary for the occurrence of a pion condensed phase with $\langle \pi^0 \rangle \neq 0$. Similar boundaries occur at the appropriate branches of

$$m_u = \frac{-m_s m_d}{\pm m_s \pm m_d} \quad (7)$$

The full phase diagram is sketched in Fig. 2.

In the CP violating phases the vacuum fluctuates about non-trivial complex matrices of form

$$\Sigma = \begin{pmatrix} e^{i\phi_1} & 0 & 0 \\ 0 & e^{i\phi_2} & 0 \\ 0 & 0 & e^{-i\phi_1 - i\phi_2} \end{pmatrix} \quad (8)$$

where the angles satisfy

$$m_u \sin(\phi_1) = m_d \sin(\phi_2) = -m_s \sin(\phi_1 + \phi_2) \quad (9)$$

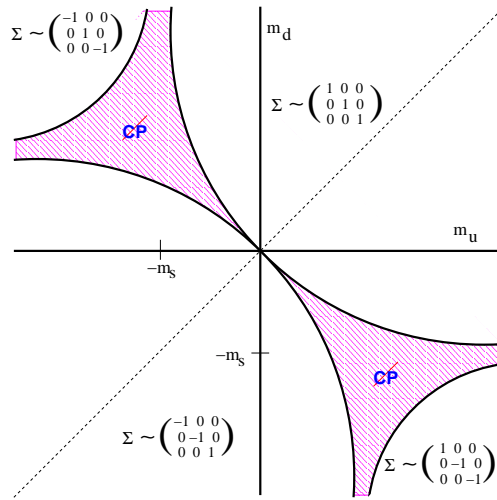


FIGURE 2. The phase diagram of QCD as a function of the up and down quark masses with a fixed positive strange quark mass.

The CP violating phases are separated from the conserving ones by second order transition lines occurring when $m_{\pi^0} = 0$. The former have two degenerate vacua related by $\phi_i \leftrightarrow -\phi_i$.

AMBIGUITIES IN THE UP QUARK MASS

I now take a section through this diagram. Phenomenologically, the down and strange quark masses appear to definitely not vanish. Consider fixing them at some positive values and study the dependence of the theory as a function of the up quark mass alone. Thus follow a horizontal line at fixed m_d in Fig. 2. To enable continuation around the boundary of the CP violating phase, extend the m_u dependence into the complex plane. This gives the qualitative structure sketched in Fig. 3.

The expectation is a first order transition along negative $\text{Re } m$ axis. This ends at a second order critical point at non-zero $\text{Re } m < 0$. Along the first order line there is a spontaneous breaking of CP. The transition has a simple order parameter $\langle \pi_0 \rangle$. The presence of the gap below $m_u = 0$ and the CP violating phase were noted some time ago by Di Vecchia and Veneziano [5].

Note that nothing significant occurs at $m_u = 0$ when $m_d \neq 0$. This raises an interesting question: Does $m_u = 0$ have any physical significance? I now argue that this is not a well posed question if $m_d \neq 0$ and $m_s \neq 0$. One consequence of this observation is that a vanishing up quark mass is an unacceptable solution to the strong CP problem.

A crucial message here is that the concept of an “underlying basic Lagrangian” does not exist. Field theory is full of divergences that must be regulated. It is only the underlying symmetries of the theory that remain significant. The case of a single massless quark gives no special symmetry because of the anomaly. Unlike the multiple degenerate quark case, no exact Goldstone bosons should appear at $m_u = 0$.

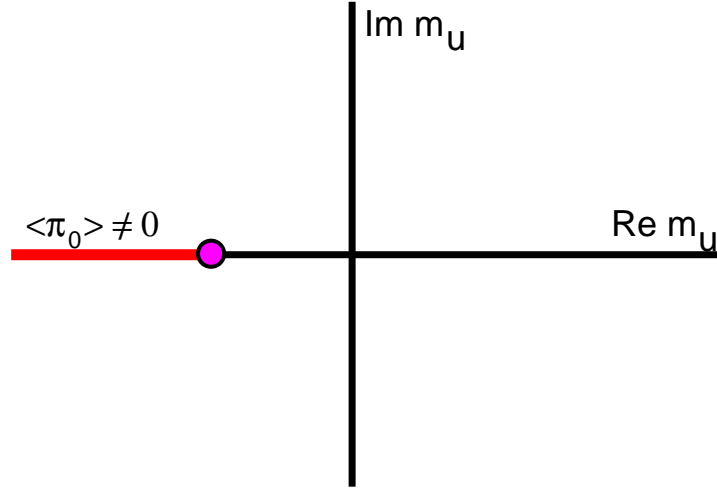


FIGURE 3. The qualitative phase diagram as a function of a complex up quark mass with fixed strange and down quark masses. (Here I ignore the reappearance of a CP conserving phase at large negative up quark mass.)

The renormalization group trajectory

A continuum field theory is defined as a limit of a cutoff theory. For QCD the bare parameters are the coupling g and quark masses m_i . These must be renormalized for the continuum limit; indeed both bare parameters renormalize to zero in a well defined way given by the renormalization group equations. For this discussion I denote the cutoff as a minimum length a . This corresponds to the inverse of a large momentum scale. To simplify the discussion I will consider a single quark mass. Then the well known flow equations in the small coupling limit take the form

$$\begin{aligned} a \frac{d}{da} g &= \beta(g) = \beta_0 g^3 + \beta_1 g^5 + \dots + \text{non-perturbative} \\ a \frac{d}{da} m &= m \gamma(g) = m(\gamma_0 g^2 + \gamma_1 g^4 + \dots) + \text{non-perturbative} \end{aligned} \quad (10)$$

The first few coefficients β_0 , β_1 , and γ_0 are scheme independent. In these equations the “non-perturbative” parts fall faster than any power of g as $g \rightarrow 0$. A crucial point, to which I will return, is that these contributions are not proportional to the quark mass.

The solution to these equations is standard

$$\begin{aligned} a &= \frac{1}{\Lambda} e^{-1/2\beta_0 g^2} g^{-\beta_1/\beta_0^2} (1 + O(g^2)) \\ m &= M g^{\gamma_0/\beta_0} (1 + O(g^2)) \end{aligned} \quad (11)$$

Rewriting shows how the coupling and mass go to zero in the continuum limit $a \rightarrow 0$

$$\begin{aligned} g^2 &\sim \frac{1}{\log(1/\Lambda a)} \rightarrow 0 \\ m &\sim M \left(\frac{1}{\log(1/\Lambda a)} \right)^{\gamma_0/2\beta_0} \rightarrow 0 \end{aligned} \quad (12)$$

The first part of this equation represents the famous phenomenon of “asymptotic freedom.”

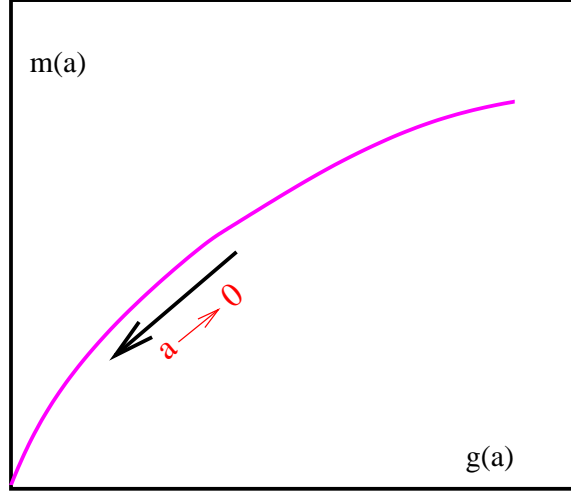


FIGURE 4. The continuum limit involves following a renormalization group trajectory in the space of bare parameters.

At a basic level these equations arise from holding a few physical quantities fixed along the “renormalization group trajectory,” as sketched in Fig. 4. For this discussion it is convenient to use the lightest baryon and the lightest meson masses, m_p , m_π , as physical quantities to hold fixed. With multiple quark flavors one would hold several meson masses fixed.

The parameters Λ and M represent “integration constants” of the renormalization group equations. Λ is conventionally interpreted as the “QCD scale,” while M defines a the “renormalized quark mass.” Their values follow from limits along the renormalization group trajectory

$$\Lambda = \lim_{a \rightarrow 0} \frac{e^{-1/2\beta_0 g^2} g^{-\beta_1/\beta_0^2}}{a} \quad (13)$$

$$M = \lim_{a \rightarrow 0} m g^{-\gamma_0/\beta_0} \quad (14)$$

The precise numerical values of Λ and M depend on the renormalization scheme.

The physical masses map directly onto the integration constants, $\Lambda = \Lambda(m_p, m_\pi)$ and $M = M(m_p, m_\pi)$. Inverting, gives the physical masses as functions of Λ and M : $\rightarrow m_i = m_i(\Lambda, M)$. Simple dimensional analysis implies this relationship must take the form $m_i = \Lambda f_i(M/\Lambda)$, with $f(x)$ some *a priori* unknown function.

In the case of multiple degenerate fermions more is known about the behavior of f . In particular, Goldstone bosons should appear as the quark mass goes to zero: $m_\pi^2 \sim m_q$. This implies the existence of a square root singularity $f_\pi(x) \sim x^{1/2}$. The location of this singularity defines what is meant by zero mass quarks, thus removing any additive ambiguity in defining M .

The single massless flavor case, however, is somewhat special. Then the meson mass $m_\pi = \Lambda f_\pi(M/\Lambda)$ does not vanish at $M = 0$. The anomaly precludes chiral symmetry and

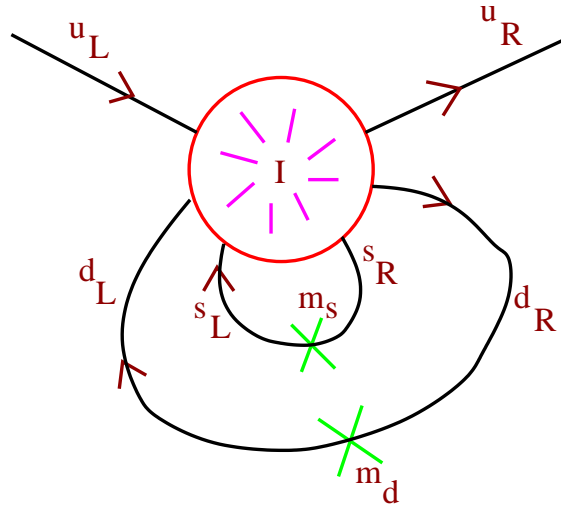


FIGURE 5. Non-perturbative effects can flip all quark spins through the anomaly. Tying the heavier quark lines together with mass terms generates a shift in the lightest quark mass.

Goldstone bosons. While the meson mass can be forced to vanish, the earlier discussion shows that this requires a special tuning to a negative quark mass. For the function $f_\pi(x)$ one expects a smooth and non-vanishing behavior at $x = 0$.

The shift of the singularity in x away from zero is due to non-perturbative effects. Indeed, non-perturbative contributions to the mass flow are expected which are not proportional to quark mass. As shown some time ago by t’Hooft [9], there are nonperturbative classical effects called “instantons” that flip all quark spins simultaneously. Tying the heavier quarks together with their mass terms as sketched in Fig. 5 gives rise to a contribution to the up quark mass flow

$$\Delta m_u \sim \frac{m_d m_s}{\Lambda_{\text{qcd}}}, \quad \Lambda_{\text{qcd}} \quad (15)$$

These effects indicate that $m_u = 0$ is not renormalization group invariant. If the up quark mass vanishes as some point on the trajectory, it will not for other points.

Matching between schemes

This non-invariance of the quark mass under the renormalization group raises the specter of scheme dependence. On changing renormalization schemes one should preserve the lowest order perturbative limit as $g \rightarrow 0$ at fixed scale a

$$\begin{aligned} \tilde{g} &= g + O(g^3) + \text{non-perturbative} \\ \tilde{m} &= m(1 + O(g^2)) + \text{non-perturbative} \end{aligned} \quad (16)$$

Here “non-perturbative” terms should vanish faster than any power of g . As mentioned above, the integration constants Λ and M in general depend on the chosen scheme.

It is important to recognize that this perturbative matching at fixed a is not the continuum limit. Indeed, taking $g \rightarrow 0$ at fixed a gives perturbation theory on free quarks, while taking $a \rightarrow 0$ at fixed g gives the divergences of field theory. For the confining physics of the real world, one must go between these limits and take a and g together to zero along the renormalization group trajectory.

To dramatically illustrate the issue, consider a particularly cooked up new scheme

$$\begin{aligned}\tilde{a} &= a \\ \tilde{g} &= g \\ \tilde{m} &= m - Mg\gamma_0/\beta_0 \times \frac{e^{-1/2\beta_0 g^2} g^{-\beta_1/\beta_0^2}}{\Lambda a}\end{aligned}\tag{17}$$

Here I have crafted the last factor in the mass expression to approach unity. This non-perturbative redefinition of the bare parameters makes

$$\tilde{M} \equiv \lim_{a \rightarrow 0} \tilde{m} \tilde{g}^{-\gamma_0/\beta_0} = M - M = 0\tag{18}$$

While this may be somewhat artificial, it shows that some scheme always exists where the renormalized quark mass vanishes. Of course doing this for the top quark will insert ridiculously large non-perturbative effects, but it is possible in principle. Because of this ambiguity, $M = 0$ is scheme dependent and thus is not a physical concept. Of course with degenerate quarks one can precisely define masslessness by the location of the square root singularity in $f(x)$ as defined above.

On the lattice

As in the general case, on the lattice the renormalization flows depend on details of the lattice action. Various gauge actions as well as the fermion formulation need to be considered. Recent discussions have concentrated on overlap/Ginsparg-Wilson fermion operators [10, 11], which bring a remnant of chiral symmetry to the lattice. However even these operators are not unique. The overlap operator relies on a projection, but it depends on the particular Dirac operator being projected. When starting with a Wilson Dirac operator, the input negative mass is adjustable over a finite range.

The one flavor theory dynamically generates a gap which will appear in the spectrum of the final Dirac operator. The overlap projection does not protect the size of this gap. With the gap present, the Ginsparg-Wilson condition is not sufficient to guarantee the preservation of $M = 0$ between schemes.

With a Ginsparg-Wilson action the concept of a massless quark is synonymous with zero topological susceptibility. If a single massless quark is an ill defined concept, this raises the question of whether the topological susceptibility is uniquely defined for $N_f < 2$. This question has also been asked in the context of making the Witten-Veneziano formula [12, 13] for the η' mass precise [14, 15]. From a perturbative point of view infinities are not a problem [16, 17]. However, to give a unique winding number to a gauge configuration requires a degree of smoothness for the gauge fields. One popular condition that ensures a well defined winding number forbids plaquettes further than

a finite distance δ from the identity [18]. However this condition is rather strong and leaves unresolved issues. In particular this “admissibility condition” has recently been shown to be in conflict with reflection positivity [19].

CONCLUSIONS

Effective chiral Lagrangians show that the strong interactions will spontaneously violate CP for large regions of parameter space. This phenomenon requires negative quark masses, a concept made physical by the anomaly. Based on this qualitative picture, I have argued that $m_u = 0$ is not a meaningful concept. As such, it cannot be regarded as a possible solution to the strong CP problem. These effects are entirely non-perturbative. As a corollary, the topological susceptibility is not uniquely defined for $N_f < 2$. Finally, I note that available simulation algorithms cannot explore this fascinating physics because it involves regions where the fermion determinant is not positive, i.e. Monte Carlo methods have a sign problem.

ACKNOWLEDGMENTS

This manuscript has been authored under contract number DE-AC02-98CH10886 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

REFERENCES

1. M. Creutz, Phys. Rev. Lett. **92**, 201601 (2004) [arXiv:hep-lat/0312018].
2. M. Creutz, Phys. Rev. Lett. **92**, 162003 (2004) [arXiv:hep-ph/0312225].
3. M. Creutz, arXiv:hep-ph/9608216.
4. R. F. Dashen, Phys. Rev. D **3**, 1879 (1971).
5. P. Di Vecchia and G. Veneziano, Nucl. Phys. B **171**, 253 (1980).
6. H. Georgi and I. N. McArthur, HUTP-81/A011
7. D. B. Kaplan and A. V. Manohar, Phys. Rev. Lett. **56**, 2004 (1986).
8. T. Banks, Y. Nir and N. Seiberg, arXiv:hep-ph/9403203.
9. G. 't Hooft, ” Phys. Rev. D **14**, 3432 (1976) [Erratum-ibid. D **18**, 2199 (1978)].
10. H. Neuberger, Phys. Lett. B **417**, 141 (1998) [arXiv:hep-lat/9707022].
11. P. H. Ginsparg and K. G. Wilson, Phys. Rev. D **25**, 2649 (1982).
12. E. Witten, Nucl. Phys. B **156** (1979) 269.
13. G. Veneziano, Nucl. Phys. B **159** (1979) 213.
14. E. Seiler and I. O. Stamatescu, MPI-PAE/PTh 10/87
15. E. Seiler, Phys. Lett. B **525** (2002) 355 [arXiv:hep-th/0111125].
16. L. Giusti, G. C. Rossi and M. Testa, Phys. Lett. B **587**, 157 (2004) [arXiv:hep-lat/0402027].
17. M. Luscher, arXiv:hep-th/0404034 (unpublished).
18. M. Luscher, Commun. Math. Phys. **85**, 39 (1982).
19. M. Creutz, Phys. Rev. D, in press (2004) [arXiv:hep-lat/0409017].